

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 5, 2002, 13:47:39 ; Search time 3997.53 seconds  
(without alignments)  
13376.994 Million cell updates/sec

Title: US-08-153-397A-1  
Sequence: 1 CGGGCCTGAGACTGGGTGA.....AAAAAAAAAACCGGATTC 3962

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database:

EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_gss:\*  
13: em\_gss\_hum:\*  
14: em\_gss\_inv:\*  
15: em\_gss\_pln:\*  
16: em\_gss\_vrl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1720	43.4	2633	11	BC006836 Mus muscu
2	960	24.2	1010	9	AL539517 AL539517
3	815.2	20.6	1019	9	AL528663 AL528663
4	815.6	20.6	987	9	AL528664 AL528664
5	761.6	19.2	898	10	BM043544 BM043544
6	715.8	18.1	770	10	BT257820 BT257820
7	715	18.0	948	10	BT257820 BT257820
8	684	17.3	799	10	BT257820 BT257820
9	680.4	17.2	715	10	BT257820 BT257820
10	672.4	17.0	744	10	BT257820 BT257820
11	672	17.0	744	10	BT257820 BT257820
12	662.6	16.7	718	9	AL043251 AL043251
13	660.2	16.7	718	9	AL043251 AL043251
14	655.6	16.5	912	10	BE304984 BE304984
15	654.6	16.5	662	10	BE249888 BE249888
16	653.8	16.5	877	10	BE249888 BE249888
17	652	16.5	957	10	BE249888 BE249888

18	651.4	16.4	653	10	BT554126 BT554126
19	650.8	16.4	683	10	BT554126 BT554126
20	650.4	16.4	685	10	BT554126 BT554126
21	646.4	16.3	787	10	BT554126 BT554126
22	646	16.3	787	10	BT554126 BT554126
23	644.2	16.3	684	10	BT554126 BT554126
24	643.6	16.2	898	10	BT554126 BT554126
25	640.8	16.2	801	10	BT554126 BT554126
26	639.6	16.1	1018	10	BT554126 BT554126
27	636.2	16.1	759	10	BT554126 BT554126
28	635	16.0	921	10	BT554126 BT554126
29	632.2	16.0	1009	10	BT554126 BT554126
30	631.2	15.9	759	10	BT554126 BT554126
31	629.2	15.9	711	10	BT554126 BT554126
32	628.4	15.9	1150	10	BT554126 BT554126
33	627.4	15.8	652	9	AL047539 AL047539
34	627.2	15.8	804	10	BT554126 BT554126
35	626.6	15.8	977	10	BT554126 BT554126
36	624	15.7	732	10	BT554126 BT554126
37	622	15.7	676	10	BT554126 BT554126
38	621	15.7	811	10	BT554126 BT554126
39	620.8	15.7	626	9	AW376875 AW376875
40	619.6	15.6	1037	10	BT554126 BT554126
41	618.2	15.6	701	10	BT554126 BT554126
42	613	15.5	680	10	BT554126 BT554126
43	603.6	15.2	649	10	BT554126 BT554126
44	602.4	15.2	1325	10	BT554126 BT554126
45	601	15.2	812	10	BT554126 BT554126

#### ALIGNMENTS

RESULT 1	BC006836	2633 bp	mRNA	linear	HMC 12-JUL-2001
LOCUS	Mus musculus	Similar to discoidin domain receptor family, member			
DEFINITION	1, clone IMAGE:3598142, mRNA.				
ACCESSION	BC006836				
VERSION	BC006836.1	GI:14711754			
KEYWORDS	house mouse,				
SOURCE	Mus musculus				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
REFERENCE	1 (bases 1 to 2633)				
AUTHORS	Strausberg, R.				
TITLE	Submitted (27-APR-2001) National Institutes of Health, Mammalian				
JOURNAL	Gene Collection (MGC), Cancer Genomics Office, National Cancer				
	Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,				
	USA				
REMARK	NIH-MGC Project URL: <a href="http://mgc.ncl.nih.gov">http://mgc.ncl.nih.gov</a>				
COMMENT	Contact: MGC help desk				
	Email: <a href="mailto:cgabs-remail.nih.gov">cgabs-remail.nih.gov</a>				
	Tissue Procurement: Jeffrey Green M.D.				
	cDNA Library Preparation: Life Technologies, Inc.				
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)				
	DNA Sequencing by: Baylor College of Medicine Human Genome				
	Sequencing Center				
	Center code: BCM-HGSC				
	Web site: <a href="http://www.hgsc.bcm.tmc.edu/cdna/">http://www.hgsc.bcm.tmc.edu/cdna/</a>				
	Contact: <a href="mailto:villalobos@bcm.tmc.edu">villalobos@bcm.tmc.edu</a>				
	Villalobos, D.K., Luna, R.A., Hale, S.M., Huylk, S., Lu, X., Garcia,				
	A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W.,				
	Muzny, D.M., Gibbs, R.A.,				
	Clone distribution: MGC clone distribution information can be found				
	through the I.M.A.G.E. Consortium/ILNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a>				
	Series: IRK Plate: 16 Row: 0 Column: 14				
	This clone was selected for full length sequencing because it				
	passed the following selection criteria: Hexamer frequency ORF				
	analysis				

FEATURES This clone has the following problem: frame shifted.

Location/Qualifiers

1. 2633

Source /organism="Mus musculus"

/db\_xref="taxon:10090"

/clone="IMAGE:3598142"

/tissue\_type="Mammary tumor, C3(1)-Tag model. Infiltrating

ductal carcinoma, 5 month old virgin mouse."

/clone\_lib="NCI\_CGAP\_Mam6"

/lab\_host="DH10B"

/note="Vector: PCMV-SPORT6"

BASE COUNT 483 a 805 c 825 g 520 t

ORIGIN

Query Match 43.4%; Score 1720; DB 11; Length 2633;  
Best Local Similarity 79.9%; Pred. No. 0;  
Matches 2200; Conservative 0; Mismatches 310; Indels 244; Gaps 5;

```

QY 277 AGAGATGCTGCCACCCCTTAGGCCGAGGAGATCAGAGCTATGAGACAGAGGCC 336
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 121 AGAGATGCTGCCACCCCTTAGGCCGAGGAGATCAGAGCTATGAGACAGAGGCC 180
QY 337 TGTCATCTTACTGCTGCTG---CTCTGTGGCAAGTGGAGATGCTGACATGAGGAC 393
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 181 TCTCATCTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
QY 394 ATTTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 453
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 241 ATTTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
QY 454 GTGACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 513
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 301 GCATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
QY 514 AGAGCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 573
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 361 AAAGCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 420
QY 574 AGTACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 633
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 421 AGTACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 480
QY 634 GGCATGCGGGGGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 693
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 481 GGCATGCGGGGGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 540
QY 694 ATGATGCGGGGGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 753
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 541 ATGATGCGGGGGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 600
QY 754 AGGACCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 813
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 601 AGGATCCGGGGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 660
QY 814 GCTTTCACCCCGGGGCTGACCGGGGCTGACCGGGGCTGACCGGGGCTGACCGGGG 873
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 661 GCTTTCACCCCGGGGCTGACCGGGGCTGACCGGGGCTGACCGGGGCTGACCGGGG 720
QY 874 GCTTTCACCCCGGGGCTGACCGGGGCTGACCGGGGCTGACCGGGGCTGACCGGGG 933
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 721 GCTTTCACCCCGGGGCTGACCGGGGCTGACCGGGGCTGACCGGGGCTGACCGGGG 730
QY 934 CTGAGGCGGTGACTCAACGACTCCACTATGACGAGACATACCTGGGCGGAGTCACT 993
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 731 -----GGGCTCACT 740
QY 994 ATGGGGGTCTGGGCGGAGCTGGAGATGTGTGGGCTGGATGATTTAGAGAGATC 1053
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 741 ATGGCGGTCTGGGCGGAGCTGGAGATGTGTGGGCTGGATGATTTAGAGAGATC 800
QY 1054 AGGAGTGGGGGTCTGGGCGGAGCTGATGATGTGGGATGGAGCAACCACTTTCCA 1113
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```

Db 801 AGGAGCTGCGGGTCTGGGCGGAGCTATGATGAGGATGAGCAATCAGAGCTTCCCA 860
QY 1114 GTGGATATGAGGAGGAGGATGAGTGAATGACCGGCTGAGGGCTTCCAGGCTATACAG 1173
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 861 CGGGCTACGGGAGTGAATGAGTGAATGAGTGAATGAGTGAATGAGTGAATGAGTGA 920
QY 1174 TCCACTGATACACATGACACACGCTGGAGAGCCGCTGCTGCTGCTGCTGCTGCTGCT 1233
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 921 TCCACTGATACACATGACACACGCTGGAGAGCCGCTTACAGGCGGGGTGAGTATGCCGT 980
QY 1234 TCCGGGTGCGCTGCCATGCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1293
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 981 TTAAGAGGTCCCGCATGCTGGGAAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1040
QY 1294 ACCTGGGGAGCCGAGAGGCGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1353
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1041 GCTTGGAGAGCCGAGAGGCGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1100
QY 1354 TTCTGAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1413
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1101 TTCTGAGTGCATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1160
QY 1414 TCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1473
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1161 TCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1208
QY 1474 GTTGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1533
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1209 GTTGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1268
QY 1534 AGCAGCCGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1593
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1269 AACAGCAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1328
QY 1594 CCAATCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1653
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1329 CCAATCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1388
QY 1654 GGTCTCTGAGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1713
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1389 GGTCTCTGAGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1448
QY 1714 TCCCTGGGAGCACTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1773
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1449 TCCCTGGGAGCACTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1508
QY 1774 AGGAGCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1833
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1509 AGGAGCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1567
QY 1834 CGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1893
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1568 ----- 1567
QY 1894 GAGGCCGCGGCGCCGCCACACCGCCTGGGCAAAACCACCAACACCGAGCCTACAGTG 1953
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1568 -----GCTGAGTG 1577
QY 1954 GGGACTATATGAGAGCTGAGAGAGCCAGGCGCGCTTTCGCCCACTGCCAACAACA 2013
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1578 GGGACTATATGAGAGCTGAGAGAGCCAGGCGCGCTTTCGCCCACTGCCAACAACA 1637
QY 2014 GGTCCGCCATATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2073
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1638 GGTCCGCCATATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1697
QY 2074 CCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2133
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1698 CCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1757
QY 2134 TCCCTGATCTGAGCTCCGCTTCAAGAGAGAGCTTGGCAGAGGAGGAGGAGGAGGAGG 2193
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1758 TCCCTGATCTGAGCTCCGCTTCAAGAGAGAGCTTGGCAGAGGAGGAGGAGGAGGAGGAG 1817
```



Db	599	CTCCCTCAGGAGTATCCAGGGGAAGCCAGTGAACCTAAACAAAGGACACATGGCA	658
QY	3160	CCTCTGCCCCCTTCCCTCCCGACAGCCCATACCTTAAATAGAGCAGTGAAGCTGCAGGT	3219
Db	659	CCTGCGCCCTTCCCTCCCGAAGCCCATACCTTAAATAGAGGACGATGAGACTGCAGGT	718
QY	3220	GGGCGGGGCCACCCAGGAGAGTGTATGGCCCTTCCTCCCTTCGAGACACATCTGCATGT	3279
Db	719	GGGCGGGGCCACCCAGGAGAGTGTATGGCCCTTCCTCCCTTCCTGACACATCTGCATGT	778
/QY	3280	CCCCCTCTGTCTTCTTCTTCTTCTAGAAAGCCCGCTGTGCGCCACCCAGCTGATCCTGTGATG	3339
Db	779	CCCCCTCTGTCTTCTTCTTCTTCTTCTAGAAAGCCCGCTGTGCGCCACCCAGCTGATCCTGTGATG	838
QY	3340	GGATCCCTCCACCCCTCCTTAGCATATCCCTTGGGGAAGGGGGAAGAAATATAGATA	3399
Db	839	GGATCCCTCCACCCCTCCTTAGCATATCCCTTGGGGAAGGGGGAAGAAATATAGATA	898
QY	3400	GACACTGGACATGGGCCATTGGAGACACCTGGGGCCCACTGGACACACATGATCTCTGGAG	3459
Db	899	GACACTGGACATGGGCCATTGGAGACACCTGGGGCCCACTGGACACACATGATCTCTGGAG	958
QY	3460	AGTGGCGTGGCCGCCACAGTCTCTCTCC	3487
Db	959	A-TGGGCTGGGCCACAGTCTCTCTCC	985

RESULT	3
LOCUS	AL528663/c
DEFINITION	AlS28663 1019 bp mRNA linear EST 13-FEB-2001 AL528663.LT.NFL001.NBC4.Homo sapiens cDNA clone CSDD001YK21.3 prime, mRNA sequence.
ACCESSION	AL528663
VERSION	AL528663.1 GI:12792156
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
Eukaryote:	Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. I.(Pases 1 to 1019)
AUTHORS	I,M.B., Gruber,C., Jesse,J. and Polayres,D.
TITLE	Full-length cdna libraries and normalization
JOURNAL	Unpublished (2001)
COMMENT	Contact : Genoscope Genoscope - Centre National de Sequenace BP 191 91006 EVRY cedex - France Email: segr@genoscope.cns.fr, web : www.genoscope.cns.fr.
FEATURES	
source	Location/Qualifiers 1..1019

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CSDD001YK21"
/clone_11b="LTI_NFL001_NBC4"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/notes="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed
by Life Technologies. Contact: Feng Liang Life
Technologies, a division of Invitrogen 9800 Medical Center
Drive Rockville, Maryland 20850, USA Fax: (1) 301 610
8311 Email: liang@lifestech.com URL:
http://fulllength.invitrogen.com"

```

Query Match	20.68;	Score 815.2;	DB 9;	Length 1019;
- Best Local Similarity	89.08;	Pred. No. 1.8e-137;		

Matches	906:	Conservative	22:	Mismatches	85:	Indels	5:	Gaps	5:
QY	2893	AGGTCATCAGAAACGGGGGGAGTTCTTCCGGGACCAGGGCCGGGAGGTACTGTGCC	2952						
Db	1016	AAGTATCTGAGAAACGGGGGGAAKTT-TTCGGGACCAGGGCCGGACGATGTACTGTGCC	958						
QY	2953	GGCCGCTCTCTGCCCCGACAGGGCCTATATAGCTGATGCTTGGTCTGAGACCCGGAGT	3012						
Db	957	GGCMGCTCTCTG-CGGAAGGGCATATATAGATGTGATCTCTGGTCTGAGACCCGGAGT	899						
QY	3013	CTGACAGGACACCCCTTTCCTCCAGCTGCATCGGTCTCTGGCAGAGATGACACTAACA	3072						
Db	898	YTGAGCAGGACACCCCTTTTCTCCRCCTCATGTGGTCTCTGGAGAGATGACTTAACA	839						
QY	3073	CGGTGTGAATCAACATCCAGCTGCCCTCCCTCAAGGAGTATCCAGGGAGCCACTG	3132						
Db	838	CGGTGTGAATCAACATCCAGCTGCCCTCCCTCAAGGAGTATCCAGGGAGCCACTG	779						
QY	3133	AACACTAAACAGAGGACACATGCGACTCTGCCCCCTCCCTCCGACAGCCCATACC	3192						
Db	778	AACACTAAACAGAGGACACATGCGACTCTGCCCCCTCCCTCCGACAGCCCATACC	719						
QY	3193	TCTAATAGAGGAGTGAAGCTG-CAGGTGGGTGGGGCCACCAGGAGCTATAGCCCT	3251						
Db	718	TCTAATAGAGGAGTGAAGCTG-CAGGTGGGTGGGGCCACCAGGAGCTATAGCCCT	659						
QY	3252	TCTCCCTCTCTGACACACACTCATATGCCCCCTCTCTCTCTCTCTCTCTAGAACCCCT	3311						
Db	658	TCTCCCTCTCTGACACACACTCATATGCCCCCTCTCTCTCTCTCTCTCTCTAGAACCCCT	599						
QY	3312	GTCCGCCACCCACGCTGTGTC-CTGTGGATGGGATCTCTTCACCTCTCTCTAGCATCCT	3370						
Db	598	GTCCGCTACCCCTGTGTGTCTGTGTGGATGGGATTTTCCTCCACCTCTCTTTTACCATCTT	539						
QY	3371	TGGGGAGGGGTGGGGAATATAGGATGTGACAGGACATAGGCCATTTGGAGCACTGG	3430						
Db	538	TGGGGAGGGGTGGGGMGATATTKGATATGATNTTGGATTTGGCCCTTGTGACTGG	479						
QY	3431	GGCCCACTGAGCAACACTGATCTCTGAGAGGTGGCTGGCCCAAGCTCTCTCCCTG	3490						
Db	478	GGCCCACTGAGCAACACTGATCTCTGAGAGGTGGCTGGCCCAAGCTCTCTCCCTG	420						
QY	3491	TCACACTGACCCCACTGCTGAGAACTGCGGGGTGAGAGGACCAAGAGAGAGAA	3550						
Db	419	TTACATATTGGACCTATTGGCTGAGAAATTTGGGGGTGAGGAGCAAGARAGAGGAA	360						
QY	3551	AATGTTCTGTGTGCTGTCTGTACTGTGTCTTCTAGCTTGGGCTTCTCTCTCTCATC	3610						
Db	359	AATGTTCTGTGTGCTGTCTGTACTGTGTCTTCTAGCTTGGGCTTCTCTCTCTCATC	300						
QY	3611	ACCTAAACACTGAGCTGGGGGTGAGCCCGCCAGCCCTGAGCACACCCCACTTCCA	3670						
Db	299	ACCTAAACACTGAGCTGGGGGTGAGCCCGCCAGCCCTGAGCACACCCCACTTCCA	240						
QY	3671	CTTGCACTCTGTACTAGAACTTCTTAAGCCTATACGTTCTGTGAGTAAATATTGG	3730						
Db	239	CTTGCACTCTGTGCTGTACTTCTTAAGCCTATACGTTTGTGAGTAAATATTGG	180						
QY	3731	GATTGGGGGGAAGAGGAGCAACGCCCCATAGCTTGGGGTGGACATCTCTAGTGA	3790						
Db	179	GATTGGGGGGAAGAGGAGCAACGCCCCATAGCTTGGGGTGGACATCTCTAGTGA	120						
QY	3791	CTGCCACATGTATTTTCTTATTAATACTTGGGGTGTGACATTTTGGGGGAGAGAC	3850						
Db	119	CTGCCACATGTATTTTCTTATTAATACTTGGGGTGTGACATTTTGGGGGAGAGAC	60						
QY	3851	AGATTTTTACACTATATATATGAGCACTTACCTTGAGGCAATTTTAAATCCCTCAGTAG	3908						
Db	59	TGATTTTACACTATATATTTGGTCTCTCTTGAGGAAATTTTAAATATATGATTTAG	2						

LOCUS AL528664 987 bp mRNA linear EST 13-FEB-2001  
 DEFINITION AL528664 LTI\_NFL001.NBC4 Homo sapiens cDNA clone CS0DD001YK21 5  
 prime, mRNA sequence.  
 ACCESSION AL528664  
 VERSION AL528664.1 GI:12792157  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 987)  
 Li, W. B., Gruber, C., Jessee, J. and Polayes, D.  
 Full-length cDNA libraries and normalization  
 JOURNAL Unpublished (2001)  
 COMMENT Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 Evry cedex - France  
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.  
 FEATURES  
 source  
 1. 987  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="CS0DD001YK21"  
 /clone\_1lb="LTI\_NFL001.NBC4"  
 /sex="male"  
 /tissue\_type="neuroblastoma cells"  
 /lab\_host="DH10B"  
 /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA  
 was primed with a NotI-oligo(dT) primer. Five prime end  
 enriched, double-stranded cDNA was digested with Not I and  
 cloned into the Not I and Eco RV sites of the pCMVSPORT 6  
 vector. Library was normalized. Library was constructed  
 by Life Technologies. Contact : Feng Liang Life  
 Technologies, a division of Invitrogen 9800 Medical Center  
 Drive Rockville, Maryland 20850, USA Fax : (1) 301 610  
 8371 Email : fliang@life.com URL :  
 http://fulllength.invitrogen.com"  
 BASE COUNT 186 a 263 c 332 g 200 t 6 others  
 ORIGIN  
 Query Match 20.6%; Score 814.6; DB 9; Length 987;  
 Best Local Similarity 98.7%; Pred. No. 2,3e-137;  
 Matches 848; Conservative 4; Mismatches 3; Indels 4; Gaps 3;  
 277 AGAGATGCTCCGCCACCCCTTAGGCCGCGGAGATCAGAGCATGAGGACAGAGGCC 336  
 |||:|||||  
 90 AGARATGCTCCGCCACCCCTTAGGCCGCGGAGATCAGAGCATGAGGACAGAGGCC 149  
 |||:|||||  
 337 TGTGATCTTATTAGTGTGCTGCTTGTGTCAGATGATGATGATGATGATGATGAT 396  
 |||:|||||  
 150 TGTGATCTTATTAGTGTGCTGCTTGTGTCAGATGATGATGATGATGATGATGAT 209  
 |||:|||||  
 397 TGTGATCTTATTAGTGTGCTGCTTGTGTCAGATGATGATGATGATGATGATGATGAT 456  
 |||:|||||  
 210 TGTGATCTTATTAGTGTGCTGCTTGTGTCAGATGATGATGATGATGATGATGATGAT 269  
 |||:|||||  
 457 ACATCTCTGCTTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 516  
 |||:|||||  
 270 ACATCTCTGCTTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 329  
 |||:|||||  
 517 GCAGTACGAGGAGTGGGCGCTGTGCGCCGAGGTCGATGTTTCCCAAGAGAGAGAGT 576  
 |||:|||||  
 330 GCAGTACGAGGAGTGGGCGCTGTGCGCCGAGGTCGATGTTTCCCAAGAGAGAGAGT 387  
 |||:|||||  
 577 ACTTGAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 636  
 |||:|||||  
 388 ACTTTAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 447  
 |||:|||||  
 637 ATGCGGGGGGCTGGGCAAGAGATTTCCCGAGATACCGGAGTGGCTTACTCCGGGAGT 696  
 |||:|||||  
 448 ATGCGGGGGGCTGGGCAAGAGATTTCCCGAGATACCGGAGTGGCTTACTCCGGGAGT 507  
 |||:|||||

697 GTGCGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 756  
 |||:|||||  
 508 GTGCGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 567  
 |||:|||||  
 757 ACCCTGAGGAGTGTGCTGTAAGAGACCTTGGGCGCCCATGGTGGCGGAGTTCGCT 816  
 |||:|||||  
 568 ACCCTGAGGAGTGTGCTGTAAGAGACCTTGGGCGCCCATGGTGGCGGAGTTCGCT 627  
 |||:|||||  
 817 TTTACCCCGGGCTGACCGGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 876  
 |||:|||||  
 628 TTTACCCCGGGCTGACCGGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 687  
 |||:|||||  
 877 TCTGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 936  
 |||:|||||  
 688 TCTGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 746  
 |||:|||||  
 937 AGCGCGTGTACTCAACGACTCCACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 996  
 |||:|||||  
 747 AGCGCGTGTACTCAACGACTCCACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 806  
 |||:|||||  
 997 GGGGTCTGGGCCAGCTGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1056  
 |||:|||||  
 807 GGGGTCTGGGCCAGCTGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 866  
 |||:|||||  
 1057 AGCTGGGGGTCTGGGCCAGCTGCTGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1116  
 |||:|||||  
 867 AGCTGGGGGTCTGGGCCAGCTGCTGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 925  
 |||:|||||  
 1117 GCTATGTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1135  
 |||:|||||  
 926 GCTATGTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 944  
 |||:|||||  
 RESULT 5  
 LOCUS BM043544 898 bp mRNA linear EST 07-NOV-2001  
 DEFINITION BM043544 60361969P1 NIH\_MGC\_40 Homo sapiens cDNA clone IMAGE:5445617 5',  
 mRNA sequence.  
 ACCESSION BM043544  
 VERSION BM043544.1 GI:16772811  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 898)  
 NIH-MGC http://mgi.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgabs-remail.nih.gov  
 Tissue Procurement: DCTD/DTF  
 cDNA Library Preparation: Ling Hong/Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.llnl.gov  
 Plate: LNCM1926 row: j column: 18  
 High quality sequence stop: 837.  
 Location/Qualifiers  
 1. 898  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5445617"  
 /clone\_1lb="NIH\_MGC\_40"  
 /tissue\_type="carcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: prostate; Vector: pORF7; Site: 1: XhoI;  
 Site: 2: EcoRI; cDNA made by oligo-dT priming.  
 Directionally cloned into EcoRI/XhoI sites using the  
 following 5' adaptor: GGACAGG(G). Library constructed by  
 Ling Hong in the laboratory of Gerald M. Rubin (University

of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies).

Note: this is a NIH-MGC library."

BASE COUNT 184 a 261 c 274 g 179 t

ORIGIN

Query Match 19.2%; Score 761.6; DB 10; Length 898;  
Best Local Similarity 96.3%; Pred. No. 8.8e-128;  
Matches 865; Conservative 0; Mismatches 24; Indels 9; Gaps 8;

```

10Y 2313 GCCAGCTTCCTCCTGCTCCAGGAATGATTTCTGAAAGAGGAGATCATGTCCAGG 2372
      |||
      2 GCCAGCTTCCTCCTGCTCCAGGAATGATTTCTGAAAGAGGAGATCATGTCCAGG 61
QY 2373 CTCAGAGACCCCAACATCATTCGGTCTGGGCGTGTGTGACAGACAGCCCTCTGC 2432
      |||
      62 CTCAGAGACCCCAACATCATTCGGTCTGGGCGTGTGTGACAGACAGCCCTCTGC 121
QY 2433 ATGATTAATGATGATGAGAAAGGCGACCTCAACAGTTCTCACTGCTCCACACCTG 2492
      |||
      122 ATGATTAATGATGATGAGAAAGGCGACCTCAACAGTTCTCACTGCTCCACACCTG 181
QY 2493 GAGAACAAAGCAGCCGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 2552
      |||
      182 GAGAACAAAGCAGCCGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 241
QY 2553 AGCTACCCATGCTGCTGATGAGACAGCCAGATGCTCCGGAGATGCTATCTGCTG 2612
      |||
      242 AGCTACCCATGCTGCTGATGAGACAGCCAGATGCTCCGGAGATGCTATCTGCTG 301
QY 2613 ACACTCACTTTGATGATGAGACAGCCAGATGCTCCGGAGATGCTATCTGCTG 2672
      |||
      302 ACACTCACTTTGATGATGAGACAGCCAGATGCTCCGGAGATGCTATCTGCTG 361
QY 2673 AACATCAATTCGACAGCTTTGGCAGAGCCGGAACCTCTATGCTGGGAGCTATTC 2732
      |||
      362 AACATCAATTCGACAGCTTTGGCAGAGCCGGAACCTCTATGCTGGGAGCTATTC 421
QY 2733 GTGACAGGCGCGGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2792
      |||
      422 GTGACAGGCGCGGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 481
QY 2793 AAGTTCAAGCTGAGTGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2852
      |||
      482 AAGTTCAAGCTGAGTGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 541
QY 2853 CTCTGTAAGGCGCCAGCCCTTTGGGAGCTACACGAGAGAGTCAATCGAAGCGGG 2912
      |||
      542 CTCTGTAAGGCGCCAGCCCTTTGGGAGCTACACGAGAGAGTCAATCGAAGCGGG 601
QY 2913 GAGTTCTTCGGGAGACGAGGCGCGGAGTGTACCTGCTCCGGGAGCTGCTCCGCG 2972
      |||
      602 GAGTTCTTCGGGAGACGAGGCGCGGAGTGTACCTGCTCCGGGAGCTGCTCCGCG 660
QY 2973 GGGCTATAT- GAGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3030
      |||
      661 GGGCTATATGAGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
QY 3031 TTTCACAGCTGCAATC- GGTTCCTGGGAG- AGGATGCAATC- AACAGGCTGTGAATACAC 3087
      |||
      721 TTTCACAGCTGCAATC- GGTTCCTGGGAG- AGGATGCAATC- AACAGGCTGTGAATACAC 780
QY 3088 ATCCAGCTGCTCCCTCCCTCAGGAGTGTATCCAGGGG- -AAGCAGTGAACACTTAAACAG 3145
      |||
      781 ATCAAGTGTGCTTCCTCAGGAGTGTATCCAGGGGAGGAGGAGGAGGAGGAGGAGG 840
QY 3146 AGGACACATGGAACCTCTG- CCCTTCCTCCCTCCGACAGCCCATCACTCTTAATAG 3202
      |||
      841 AGGACACATGGAACCTCTG- CCCTTCCTCCCTCCGACAGCCCATCACTCTTAATAG 898

```

RESULT 6

BI257820  
LOCUS BI257820 770 bp mRNA linear EST 17-JUL-2001  
DEFINITION 602969918F1 NIH\_MGC\_12 Homo sapiens cDNA clone IMAGE:5109436 5',  
mRNA sequence.  
ACCESSION BI257820  
VERSION BI257820.1 GI:14813571  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 770)  
NIH-MGC <http://mgi.nci.nih.gov/>,  
National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
CONTACT: Robert Strausberg, Ph.D.  
Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
Tissue Procurement: ATCC  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: Incyte Genomics, Inc.  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LINL at:  
<http://image.llnl.gov>  
Plate: L14M1265 row: c column: 05  
High quality sequence stop: 765.  
Location/Qualifiers  
1..770  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_1="IMAGE:5109436"  
/clone\_1b="NIH\_MGC\_12"  
/tissue\_type="cervical carcinoma cell line"  
/lab\_host="DH10B"  
/note="Organ: cervix; Vector: pCMV-Sport6; Site: 1: NotI;  
Site: 2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 1.4 kb. Library prepared by Life  
Technologies."

#### FEATURES

source  
1..770  
Location/Qualifiers

BASE COUNT 144 a 207 c 262 g 157 t

Query Match 18.1%; Score 715.8; DB 10; Length 770;  
Best Local Similarity 97.3%; Pred. No. 1.7e-119;  
Matches 750; Conservative 0; Mismatches 17; Indels 4; Gaps 2;

```

QY 229 GAAGGTGGCTATTCATGACGATGGGGTGTGACTTGAAGAAATGCCAAGATGCTGCC 288
      |||
      1 GAAGGTGGCTATTCATGACGATGGGGTGTGACTTGAAGAAATGCCAAGATGCTGCC 60
QY 289 CCCACCCCTTAGCCCGAGAGGATGACGATGAGGAGCCAGAGCCCTGCTCATCTTAC 348
      |||
      61 CCCACCCCTTAGCCCGAGAGGATGACGATGAGGAGCCAGAGCCCTGCTCATCTTAC 120
QY 349 TGCTGCTGCTCTTGTGCAAGTGAATGCTGACATGAAGGACATTTTGTATCTGCA 408
      |||
      121 TGCTGCTGCTCTTGTGCAAGTGAATGCTGACATGAAGGACATTTTGTATCTGCA 180
QY 409 AGTCCCGCTATGCGCTGGGAGTCAAGACCGACATCCAGACAGATCTGCTT 468
      |||
      181 AGTCCCGCTATGCGCTGGGAGTCAAGACCGACATCCAGACAGATCTGCTT 240
QY 469 CCAGCTCTGCTGATGATTCACATGACCGCCCGACAGAGAGTGAAGAGCAGTGAAGG 528
      |||
      241 CCAGCTCTGCTGATGATTCACATGACCGCCCGACAGAGAGTGAAGAGCAGTGAAGG 300
QY 529 ATGGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 588
      |||
      301 ATGGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
QY 589 ATCTACAAGACATCCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 648
      |||
      361 ATCTACAAGACATCCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420

```



```

OY 649 TGGGCAAGAGATTCCTCCGGAGCTACCGGCTGCTACTCCCGGAGTGGCGCGTGA 708
DB 421 TGGGCAAGAGATTCCTCCGGAGCTACCGGCTGCTACTCCCGGAGTGGCGCGTGA 480
OY 709 TGGGCTGGAAGACCGCTGGGGTCAAGAGATGATCTCAGGCAATGAGACCCGAGAGAG 768
DB 481 TGGGCTGGAAGACCGCTGGGGTCAAGAGATGATCTCAGGCAATGAGACCCGAGAGAG 540
OY 769 TGGTCTGAAGACCTTGGGCCCCCATGCTGTCGCCAGTGGTGGCTTACCCCGGG 828
DB 541 TGGTCTGAAGACCTTGGGCCCCCATGCTGTCGCCAGTGGTGGCTTACCCCGGG 600
OY 829 CTGACCGGGTCAATG---AGTGTCTGTCTGCGGGTAGAGCTATGAGTGGCTCCCTGAGAG 885
DB 601 CTGACCGGGTCAATGAGCACTGTACTGCGGGTAGAGCTATGAGTGGCTCCCTGAGAG 660
OY 886 ATGAGCTCCTGTCTTACACCGGCCCCGTGGGGCAGACAAATGATATGATGAGAGCGGTG 945
DB 661 ATGAGCTCCTGTCTTACACCGGCCCCGTGGGGCAGACAAATGATATGAGAGCGGTG 719
OY 946 ACCTAAGAGCTCCACTATGACGACATACCTGGGGCGGAGTCAATGATG 996
DB 720 ACCTAAGAGCTCCACTATGACGACATACCTGGGGCGGAGTCAATGATG 770

RESULT 7
BI825684 948 bp mRNA linear EST 04-OCT-2001
LOCUS 60307266Bf1 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5164480 5'
DEFINITION mRNA sequence.
ACCESSION BI825684
VERSION BI825684.1 GI:15937234
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 948)
NIH-MGC http://mgc.ncl.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: rgs@ds-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
DNA sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LHM11408 Row: h Column: 17
High quality sequence stop: 785.
FEATURES
Source
1. 948
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5164480"
/clone_id="NIH_MGC_119"
/tissue_type="medulla"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMV-Sport6; Site:1: NciI;
Site:2: EcoRV (destroyed); RNA source normal medulla from
anonymous male age 27. Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.3 kb, insert size range
0.9-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 013. Note:
this is a NIH-MGC Library."
BASE COUNT 178 a 272 c 296 g 202 t
ORIGIN

```

```

Query Match 18.0%; Score 715; DB 10; Length 948;
Best Local Similarity 93.0%; Pred. No. 2,3e-119;
Matches 845; Conservative 0; Mismatches 35; Indels 29; Gaps 8;

OY 2136 CCTCGATCTCGATCCCGCTTCAAGAGAAAGCTTGGCGAGGCCCACTTTGGAGAGTGCAC 2195
DB 1 CCTCGATCTCGATCCCGCTTCAAGAGAAAGCTTGGCGAGGCCCACTTTGGAGAGTGCAC 60
OY 2196 CTGTGTGAGGTGACACAGCCCTCAAGATCTGTGACGCTTGTGATTTCCCTTAATGGGT 2255
DB 61 CTGTGTGAGGTGACACAGCCCTCAAGATCTGTGACGCTTGTGATTTCCCTTAATGGGT 120
OY 2256 AAGGACACCTTTGCTGTGAGTCTCAAGATCTTACGCCACAGATCCACCAAGATGCC 2315
DB 121 AAGGACACCTTTGCTGTGAGTCTCAAGATCTTACGCCACAGATCCACCAAGATGCC 178
OY 2316 AGCTTCTCCTTGTCTCCAGGAAATGATTTCTGAAAGAGGTGAAGATCATGTGAGGCTC 2375
DB 179 -----CCAGGAATGATTTCTGAAAGAGGTGAAGATCATGTGAGGCTC 222
OY 2376 AAGGACCCCAATCATTTGGGTGCTGGGCGTGTGTGTCAGAGAGAGACCCCTTGCATG 2435
DB 223 AAGGACCCCAATCATTTGGGTGCTGGGCGTGTGTGTCAGAGAGAGACCCCTTGCATG 282
OY 2436 ATTACTGACTATGAGAGAACGGGCACTCAACAGATTCTCAGTGCACACAGCTGGAG 2495
DB 283 ATTACTGACTATGAGAGAACGGGCACTCAACAGATTCTCAGTGCACACAGCTGGAG 342
OY 2496 GACAAAGGACCGGAGAGGGGCCCCCTGGGACGGGACAGGCTGCGCAGAGGCCACCATCAGC 2555
DB 343 GACAAAGGACCGGAGAGGGGCCCCCTGGGACGGGACAGGCTGCGCAGAGGCCACCATCAGC 402
OY 2556 TACCAATGCTGCTCATGTGAGGACCCAGATGCGCTCCGGAGTGGCTATCTGGCCACA 2615
DB 403 TACCAATGCTGCTCATGTGAGGACCCAGATGCGCTCCGGAGTGGCTATCTGGCCACA 462
OY 2616 CTCAACTTTGTACATCGGAGACCTGCGCACAGCGGAACCTGATTTGGGAAAAATTTAC 2675
DB 463 CTCAACTTTGTACATCGGAGACCTGCGCACAGCGGAACCTGATTTGGGAAAAATTTAC 522
OY 2676 ATCAAAATCGCAGACTTTGGCATGAGCCGGAACCTTATGCTGGGACTATTTACCTGTG 2735
DB 523 ATCAAAATCGCAGACTTTGGCATGAGCCGGAACCTTATGCTGGGACTATTTACCTGTG 582
OY 2736 CAGGCGCGGCGAGTCTGCCATCCGCTGGAGTGGCTGGAGTGCATCTATGGGGAAG 2755
DB 583 CAGGCGCGGCGAGTCTGCCATCCGCTGGAGTGGCTGGAGTGCATCTATGGGGAAG 642
OY 2796 TTCAGAGCTGGAGAGTGAAGTGTGGGCTTTGGTGTGA--CCCTGAGGAGTGTGATGCT 2854
DB 643 TTCAGAGCTGGAGAGTGAAGTGTGGGCTTTGGTGTGA--CCCTGAGGAGTGTGATGCT 702
OY 2855 CTGTAGGCGCCACAGCCCTTTGGG---CAGCTCACGAGCAGAGTCAATGAGAAAGCGGG 2911
DB 703 CTGTAGGCGCCACAGCTTTGGGCGAGCTCAACCGAGCAGAGTCAATGAGAAAGCGGG 762
OY 2912 GGAGTTCTTCGGGACGAGGCGCGGCA--GGTGTACCTGCTCCCGGCC--GGCTGCTGCGC 2967
DB 763 GGAGTTCTTCGGGACGAGGCGCGGCA--GGTGTACCTGCTCCCGGCC--GGCTGCTGCGC 822
OY 2968 CGCAGGCGCTA--TATGAGCTGATGCTTGTGCTGAGACCGGAGTGTGACAGACGACA 3026
DB 823 CGCAGGCGCTA--TATGAGCTGATGCTTGTGCTGAGACCGGAGTGTGACAGACGACA 880
OY 3027 CCCCTTTTC 3035
DB 881 CCTTTTCC 889

RESULT 8
BG912321 799 bp mRNA linear EST 05-JUN-2001
LOCUS

```

DEFINITION 602806864F1 NCI-CGAP\_Brn67 Homo sapiens cDNA clone IMAGE:4939239  
5', mRNA sequence.  
ACCESSION BC912321  
VERSION BC912321.1 GI:14292797  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 799)  
REFERENCE NIH-MGC http://mgi.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
AUTHORS Contact: Robert Strausberg, Ph.D.  
Email: cgaaps-remail.nih.gov  
Tissue Procurement: David N. Louis, M.D.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Plate: LHAM0875 row: o column: 16  
High quality sequence stop: 768.  
Location/Qualifiers  
1. 799  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4939239"  
/clone\_lib="NCI-CGAP\_Brn67"  
/tissue\_type="anaplastic oligodendroglioma with 1p/19q  
loss"  
/lab\_host="DH10B (r1 phage-resistant)"  
/note="Organ: brain; Vector: PCMV-SPORE; Site: 1; Note:  
Site: 2; Salt: Cloned unidirectionally. Primer: oligo dt.  
Average insert size 2.3 kb. Constructed by Life  
Technologies. Note: this is a NCI-CGAP library."  
BASE COUNT 163 a 238 c 251 g 147 t  
ORIGIN

Query Match 17.3% Score 684 DB 10 Length 799;  
Best Local Similarity 95.7% Pred. NO. 9.4e-114;  
Matches 758; Conservative 0; Mismatches 25; Indels 9; Gaps 5;

QY 2443 ACTACATGAGAACGGGACCTCAACCACTCTCACTGCGCCACAGTGGAGACAGG 2502  
Db 1 ACTACATGAGAACGGGACCTCAACCACTCTCACTGCGCCACAGTGGAGACAGG 60  
QY 2503 CAGCCGAGGGGGCCCCCTGGGGAGCGGAGGCTGGCAGGGGCCACCATCAGTACCCAA 2562  
Db 61 CAGCCGAGGGGGCCCCCTGGGGAGCGGAGGCTGGCAGGGGCCACCATCAGTACCCAA 120  
QY 2563 TGCCTGCGCATGTGGACAGCCAGATGCGCTCCCGGCACTGCTATCTGGCCACTCACT 2622  
Db 121 TGCCTGCGCATGTGGACAGCCAGATGCGCTCCCGGCACTGCTATCTGGCCACTCACT 180  
QY 2623 TTGTACATCGGGAGCTGGCCACGCGGAACTGCTAGTGGGAAATTTACATCACTAAA 2682  
Db 181 TTGTACATCGGGAGCTGGCCACGCGGAACTGCTAGTGGGAAATTTACATCACTAAA 240  
QY 2683 TCGCAGACTTTGGCATGAGCCGGAACCTCTATGCTGGGGACTATTACCGTGTGAGGGCC 2742  
Db 241 TCGCAGACTTTGGCATGAGCCGGAACCTCTATGCTGGGGACTATTACCGTGTGAGGGCC 300  
QY 2743 GGGCAGTGTCTCCCATCGGATGGCTGGGAGTGCATCTTCATGGGGAAGTTCACGA 2802  
Db 301 GGGCAGTGTCTCCCATCGGATGGCTGGGAGTGCATCTTCATGGGGAAGTTCACGA 360  
QY 2803 CTGGAGTGAAGTGTGGGCTTTGGTGTGACCTGTGTGGAGTGTCTGATGCTGTAGGG 2862  
Db 361 CTGGAGTGAAGTGTGGGCTTTGGTGTGACCTGTGTGGAGTGTCTGATGCTGTAGGG 420

QY 2863 CCCAGCCCTTTGGGACAGCTACCGACGAGAGTGCATCGAAGACGGGGAGTTCTCC 2922  
Db 421 CCCAGCCCTTTGGGACAGCTACCGACGAGAGTGCATCGAAGACGGGGAGTTCTCC 480  
QY 2923 GGGACACAGGGCCGGACAGGTGTACCTGTCCGGCCGCTGCTGCCGC -AGGGCTATAT 2981  
Db 481 GGGACACAGGGCCGGACAGGTGTACCTGTCCGGCCGCTGCTGCCGCAGAGGCTATAT 540  
QY 2982 GAGCTGATGCTTGGCTGCTGCTGAGCCGGAGTGTGAGAGGACACACCTTTCCAGACTG 3041  
Db 541 GAGCTGATGCTTGGCTGCTGCTGAGCCGGAGTGTGAGAGGACACCA -CCTTTCCAGACTG 599  
QY 3042 CATCGGTCTCTGGCAGAGATGACATCAACAGGCTGTGAATCAACATCCAGCTG----- 3096  
Db 600 CATCGGTCTCTGGCAGAGATGACATCAACAGGCTGTGAATCAACATCCAGCTGCGC 659  
QY 3097 CCCCTCCCTTACAGGAGTGTATCCAGGGAGAGCC -AGTACACTAA -CAAGAGACACAA 3154  
Db 660 CCCCTCCCTTACAGGAGTGTATCCAGGGAGAGCCAAAGTACACTAAACCAAGAGAGACAA 719  
QY 3155 TGGCAGCTGTGCTCCCTCCCTCCCGACAGCCCATCAGCTCTAATAGAGGAGTGAAGT 3214  
Db 720 TGGCAGCTGTGCTCCCTCCCTCCCGACAGCCCATCAGCTCTAATAGAGGAGTGAAGTGA 779  
QY 3215 CAGGTGGGCTGG 3226  
Db 780 CAGGTGGGCTGGG 791

RESULT 9  
LOCUS BE744806 715 bp mRNA linear EST 15-SEP-2000  
DEFINITION 601575025F1 NIH\_MGC\_9 Homo sapiens cDNA clone IMAGE:383604 5',  
mRNA sequence.  
ACCESSION BE744806  
VERSION BE744806.1 GI:10158798  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 715)  
REFERENCE NIH-MGC http://mgi.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
AUTHORS Contact: Robert Strausberg, Ph.D.  
Email: cgaaps-remail.nih.gov  
Tissue Procurement: DCTD/DTF  
CDNA Library Preparation: Ling Hong/Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at: image.llnl.gov  
Plate: LHAM519 row: a column: 07  
High quality sequence stop: 713.  
Location/Qualifiers  
1. 715  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3836046"  
/clone\_lib="NIH\_MGC\_9"  
/tissue\_type="adenocarcinoma cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: ovary; Vector: pORF7; Site: 1; Note:  
ECORI: cDNA made by oligo-dt priming. Directionally  
cloned into EcoRI/XhoI sites using the following 5'  
adaptor: GGACGAG(G). Size selected >500bp for average  
insert size 1.8kb. Library constructed by Ling Hong in  
the laboratory of Gerald M. Rubin (University of  
California, Berkeley) using Zap-cDNA synthesis kit  
(Stratagene) and superscript II RT (Life Technologies)."  
BASE COUNT 130 a 226 c 212 g 147 t  
ORIGIN





```

OY 3057 GAGATGCACTCAACA-CGGTGTGAATCACAATCAGCTCCCTCCCTCAGGAGTGA 3115
DB 659 GAGATGCACTCAACAACCGGTGTGAATCACAATCAG-TGGCCCTCCCTCAGGAGTGA 717
OY 3116 TC 3117
DB 718 TC 719

RESULT 11
LOCUS B1193181 744 bp mRNA linear EST 10-JUL-2001
DEFINITION 602947157F1 NIH_MGC_42 Homo sapiens cDNA clone IMAGE:5090223 5',
mRNA sequence.
ACCESSION B1193181
VERSION B1193181.1 GI:14648201
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 744)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: L1CM1852 row: b. column: 16
High quality sequence stop: 675.
Location/Qualifiers
1. 744
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="5090223"
/clone_lib="NIH_MGC_42"
/tissue_type="epitheloid carcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: pancreas; Vector: pCMV7; Site_1: XhoI;
Site_2: EcoRI; CDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCAGAG(g). Size-selected >500bp
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library.]"
BASE COUNT 149 a 212 c 229 g 154 t
ORIGIN
Query Match 17 0%; Score 672; DB 10; Length 744;
Best Local Similarity 97.5%; Pred. No. 1,4e-111;
Matches 714; Conservative 0; Mismatches 15; Indels 3; Gaps 3;
OY 2237 TTTCCCTTAATGTGCGTAAGGACACCTTGTGTGATCAAGATCTTACGGCC 2296
DB 2 TTTCCCTTAATGTGCGTAAGGACACCTTGTGTGATCAAGATCTTACGGCC 61
OY 2297 AGATGCCACCAAGATGCCAGCTTCTCTGTCAGGAATGATTCTCGAAGAAGGT 2356
DB 62 AGATGCCACCAAGATGCCAGCTTCTCTGTCAGGAATGATTCTCGAAGAAGGT 121
OY 2357 GAAGATCATGTGAGGCTCAAGAGACCCCAATCATCTGGCTCTGGGCTGTGTGCA 2416
DB 122 GAAGATCATGTGAGGCTCAAGAGACCCCAATCATCTGGCTCTGGGCTGTGTGCA 181

```

```

OY 2417 GGAGACCCCTCTGCATGATTACTGACTACAGAGAGGCGACCTCAACAGTTGCT 2476
DB 182 GGAGACCCCTCTGCATGATTACTGACTACATGAGAGGCGACCTCAACAGTTGCT 241
OY 2477 CAGTCCACCAAGCTGAGAGCAAGGACGCGAGGCGCCCTGGGAGCGGAGCTGC 2536
DB 242 CAGTCCACCAAGCTGAGAGCAAGGACGCGAGGCGCCCTGGGAGCGGAGCTGC 301
OY 2537 GCAGGCGCCCAACCATGAGTACCTGCTGCTGATGTGCGAGCCAGATCGCTCCGG 2536
DB 302 GCAGGCGCCCAACCATGAGTACCTGCTGCTGATGTGCGAGCCAGATCGCTCCGG 361
OY 2597 CATCGCTATGTGCGCCACACTCACTTGTATCATGCGGAGCTGGCGACCGGAACTGCT 2656
DB 362 CATCGCTATGTGCGCCACACTCACTTGTATCATGCGGAGCTGGCGACCGGAACTGCT 421
OY 2657 AGTTGGGAAAATTTCACATCAAAATCGAGACTTTGGCATGAGCCGGAACCTCTATGC 2716
DB 422 AGTTGGGAAAATTTCACATCAAAATCGAGACTTTGGCATGAGCCGGAACCTCTATGC 481
OY 2717 TGGGACTATTACCGTGTGCAGGCGCGGAGAGTCCCATCCGCTGATGCTCGGA 2776
DB 482 TGGGACTATTACCGTGTGCAGGCGCGGAGAGTCCCATCCGCTGATGAG-CTGGGA 540
OY 2777 GTGCATCCTCATGGGAGAGTTCAGACTGCGAGTGGAGTGGGCTTTGGTGTGACCT 2836
DB 541 GTGCATCCTCATGGGAGAGTTCAGACTGCGAGTGGAGTGGGCTTTGGTGTGACCT 600
OY 2837 GTGGAGAGTGTGATGCTGTGTAGGCGCCAGCCCTTTGGGAGCTCAACGAGAGAGT 2896
DB 601 GTGGAGAGTGTGATGCTGTGTAGGAGCAAG-CTTTGGGAGAGTCAACGAGAGAGT 659
OY 2897 CATGAGAAAG-CGGGAGAGTCTTCCGGGACAGGCGGAGGTGATCTGTCCGGC 2955
DB 660 CATGAGAAAGCGCGGGAGAGTCTTCCGGGACAGGCGGAGGTGATCTGTCCGGC 719
OY 2956 CGCTGCTGCTGC 2967
DB 720 CGCTGCTGCTGC 731

RESULT 12
LOCUS BG696171 718 bp mRNA linear EST 07-MAY-2001
DEFINITION 602659285F1 NCL_CGAP_Skn3 Homo sapiens cDNA clone IMAGE:4802421 5',
mRNA sequence.
ACCESSION BG696171
VERSION BG696171.1 GI:13961044
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 718)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: James Cleaver, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA
Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: L1AM10696 row: j column: 22
High quality sequence stop: 717.
Location/Qualifiers
1. 718
/organism="Homo sapiens"
/db_xref="taxon:9606"
FEATURES
source

```

/clone-IMAGE:4802421"  
 /clone.lib="NCI\_CGAP\_Skn3"  
 /lab.host="DH10B (T1 phage-resistant)"  
 /note="Organ: skin; Vector: pCMV-SORT; Site:1: NotI;  
 Site:2: SalI; Cloned unidirectionally. Primer: Oligo dt.  
 Average insert size 1.5kb. Library constructed by Life  
 Technologies. Note: this is a NCI\_CGAP Library."  
 BASE COUNT 133 a 228 c 210 g 147 t  
 ORIGIN

Query Match 16.7% Score 662.6; DB 10; Length 718;  
 Best Local Similarity 98.7%; Pred. No. 6.9e-110;  
 Matches 710; Conservative 0; Mismatches 4; Indels 5; Gaps 4;

QY 2670 TTCACATCAAAATGCGACACTTGGCAGAGCGGACACCTCTATCTGGGAGCTATTAC 2729  
 DB 1 TTCACATCAAAATGCGACACTTGGCAGAGCGGACACCTCTATCTGGGAGCTATTAC 60  
 QY 2730 CGTGTGACAGGCGGCGAGCTGTCGCCATCCGCTGGATGGCGGGAGTGCATCCCTCAG 2789  
 DB 61 CGTGTGCA-GCGCGGAGCTGCTGCCATCCGCTGGATGGCGGGAGTGCATCCCTCAG 119  
 QY 2790 GGGAGTTTCACAGCTGCGAGTACGCTGTGGCTTGTGTGACCTGTGGAGTGTG 2849  
 DB 120 GGGAGTTTCACAGCTGCGAGTACGCTGTGGCTTGTGTGACCTGTGGAGTGTG 179  
 QY 2850 ATGCTCTGTAGAGGCCAGCCCTTTGGGCAAGCTCACCAGACAGAGTCTATCGAAGAGCG 2909  
 DB 180 ATGCTCTGTAGAGGCCAGCCCTTTGGGCAAGCTCACCAGACAGAGTCTATCGAAGAGCG 239  
 QY 2910 GGGAGTTCTTCGCGGAGACAGGCGGCGAGTGTACCTTCCCGCGCGCTCCCTGCGCG 2969  
 DB 240 GGGAGTTCTTCGCGGAGACAGGCGGCGAGTGTGTACCTTCCCGCGCGCTCCCTGCGCG 299  
 QY 2970 CAGGCGCTATATGAGCTGATGCTGTGCTGAGACCGGAGTGTGAGCAGGACACCC 3029  
 DB 300 CAGGCGCTATATGAGCTGATGCTGTGCTGAGACCGGAGTGTGAGCAGGACACCC 359  
 QY 3030 TTTTCCAGCTGACATCGGCTCTGCGAGAGATGCACTCAACAGGTTGATACACAT 3089  
 DB 360 TTTTCCAGCTGACATCGGCTCTGCGAGAGATGCACTCAACAGGTTGATACACAT 419  
 QY 3090 CCAGCTGCGCCCTCCCTCAAGGAGTATCAGGAGGAAAGCAGAGACACTAAACAGAGAG 3149  
 DB 420 CCAGCTGCGCCCTCCCTCAAGGAGTATCAGGAGGAAAGCAGAGACTAAACAGAGAG 479  
 QY 3150 CACAATGACACTTGCCTTCCCTCCAGACAGCCATCACTTAATAGAGCAGTGA 3209  
 DB 480 CACAATGACACTTGCCTTCCCTCCAGACAGCCATCACTTAATAGAGCAGTGA 539  
 QY 3210 GACTGCAGTGGGCTGGGCGCCAGGAGAGTATGCCCTTCT--CCCTTCTGTGAC 3267  
 DB 540 GACTGCAGTGGGCTGGGCGCCAGGAGAGTATGCCCTTCTCCCTTCTGTGAC 599  
 QY 3268 ACACCTCATGTCCCTT-CCGTCTCTCTCTCTAGAGAGCCCGTGGCCCAACCAAGT 3326  
 DB 600 ACACCTCATGTCCCTTACT 659  
 QY 3327 GGTCTGTGATGGAGATCTCTCAACCTCTCTAGCC-ATCCCTTGGGAGAGGCTGG 3384  
 DB 660 GGTCTGTGATGGAGATCT 718

RESULT 13  
 AL043251 772 bp mRNA linear EST 29-FEB-2000  
 LOCUS DEFZP434K0223-r1.434 (synonym: htes3) Homo sapiens cDNA clone  
 DEFINITION DEFZP434K0223 5', mRNA sequence.  
 ACCESSION AL043251  
 VERSION AL043251.1 GI:5422649  
 KEYWORDS EST.  
 SOURCE human.

ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 REFERENCE 1 (bases 1 to 772)  
 AUTHORS Blum, H., Bauersachs, S., Mewes, H.W., Gassenhuber, J. and Wiemann, S.  
 TITLE EST (Blum, et al.)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Blum H

MIPS  
 Am Klopferplatz 18a D-82152 Martinsried, Germany  
 This is the 5' sequence of the clone insert.  
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
 Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;  
 sequenced by LMU (Ludwig Maximilians University,  
 Munich/Germany) within the cDNA sequencing consortium of the German  
 Genome Project.  
 No sl sequence available.  
 This clone (DEFZP434K0223) is available at the RZPD in Berlin.  
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.  
 FEATURES  
 source location/Qualifiers  
 1..772  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="DEFZP434K0223"  
 /clone.lib="434 (synonym: htes3)"  
 /tissue\_type="testis"  
 /dev\_stage="adult"  
 /lab\_host="DH10B"  
 /note="Vector: pSPORT1; Site\_1: NotI; Site\_2: SalI"

BASE COUNT 118 a 250 c 260 g 143 t 1 others  
 ORIGIN

Query Match 16.7% Score 660.2; DB 9; Length 772;  
 Best Local Similarity 92.4%; Pred. No. 1.9e-109;  
 Matches 750; Conservative 0; Mismatches 4; Indels 58; Gaps 3;

QY 64 CTCTGGGAGCCGCTCCCGACACCCAGCCCGCGCGCTCCCGCTCCCGCTCCCG 123  
 DB 1 CTCTGGGAGCCGCTCCCGACACCCAGCCCGCGCGCTCCCGCTCCCGCTCCCG 60  
 QY 124 GCTCTGGCTCCCTCCGCTCCCGCTCCCGCTCCCGCTCCCGCTCCCGCTCCCG 183  
 DB 61 GCTCTGGCTCCCTCCGCTCCCGCTCCCGCTCCCGCTCCCGCTCCCGCTCCCG 120  
 QY 184 GGTGCGAGCGCTGGGTGCGCGGAGAGAGATGATGATGATGATGATGATGATGATGAT 243  
 DB 121 GGTGCGAGCGCTGGGTGCGCGGAGAGAGATGATGATGATGATGATGATGATGATGAT 157  
 QY 244 CTGACGATGGGTTGACATTGCAAGAGATGCTGCCACCCACCCCTTATAGC 303  
 DB 158 -----AAGATGCTGCCGCCACCCCTTATAGC 184  
 QY 304 CCGAGGATCAGAGACTATGAGACAGAGCCCTCTCATCTTACTGCTGCTCTTGG 363  
 DB 185 CCGAGGATCAGAGACTATGAGACAGAGCCCTCTCATCTTACTGCTGCTCTTGG 244  
 QY 364 TGGCAAGTGAAGTCTACATGAAGGACATTTTGATCTGCGCAAGAGCGGCTATGCC 423  
 DB 245 TGGCAAGTGAAGTCTACATGAAGGACATTTTGATCTGCGCAAGAGCGGCTATGCC 304  
 QY 424 TGGCATGACAGACCGGACATCCAGACAGTATGATCTGCTTCCAGCTCTGTGAC 483  
 DB 305 TGGCATGACAGACCGGACATCCAGACAGTATGATCTGCTTCCAGCTCTGTGAC 364  
 QY 484 ATTCCACTGCGCGCCGACAGAGAGTGGAGAGAGTGGAGAGAGAGAGAGAGAGAGAGAG 543  
 DB 365 ATTCCACTGCGCGCCGACAGAGAGTGGAGAGAGTGGAGAGAGAGAGAGAGAGAGAGAG 424  
 QY 544 CCGCAGGCTGGGTCTTCCCAAGAGAGAGAGTGGAGAGAGTGGAGAGAGAGAGAGAGAGAG 603  
 DB 425 CCGCAGGCTGGGTCTTCCCAAGAGAGAGAGTGGAGAGAGTGGAGAGAGAGAGAGAGAGAG 484



JOURNAL  
MEDLINE  
COMMENT

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
20202663  
Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome  
Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=PM0642-PM0-NN1171-  
181000-001-c05&f3=2000-10-18&f4=1)  
Seq primer: puc 18 forward  
High quality sequence/stop: 660.  
Location/Qualifiers

FEATURES  
source

1. 662  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="NN1171"  
/dev\_stage="Adult"  
/note="Organ: nervous\_normal; Vector: puc18; Site\_1: SmaI;  
Site\_2: SmaI; A mini-library was made by cloning products  
derived from ORESTES PCR (U.S. Letters Patent application  
No. 196,716 - Ludwig Institute for Cancer Research)  
profiles into the puc 18 vector. Reverse transcription of  
tissue mRNA and cDNA amplification were performed under  
low stringency conditions."

BASE COUNT 124 a 214 c 191 g 132 t 1 others  
ORIGIN

Query Match 16.58; Score 654.6; DB 10; Length 662;  
Best Local Similarity 99.28; Pred. No. 1.9e-108;  
Matches 657; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2798 CACGACTGCGAGTACGAGTGGGCTTGGTGTGACCTGTGGAGGAGTGCATGCTCTG 2857  
DB 1 CACGACTGCGAGTACGAGTGGGCTTGGTGTGACCTGTGGAGGAGTGCATGCTCTG 60  
QY 2858 TAGGGCCAGCCCTTTGGGAGCTCACCAGAGAGAGTCAATCAGAAAGCGGGAGTT 2917  
DB 61 TAGGGCCAGCCCTTTGGGAGCTCACCAGAGAGAGTCAATCAGAAAGCGGGAGTT 120  
QY 2918 CTTCCGGAGCAGGCGGCGAGGTGACCTGCTCCGGCCCTGCTGCCGAGGGCT 2977  
DB 121 CTTCCGGAGCAGGCGGCGAGGTGACCTGCTCCGGCCCTGCTGCCGAGGGCT 180  
QY 2978 ATATGAGCTGATGCTTGGGCTGAGCGGAGTCTGAGAGGAGCAGCCCTTTCCCA 3037  
DB 181 ATATGAGCTGATGCTTGGGCTGAGCGGAGTCTGAGAGGAGCAGCCCTTTCCCA 240  
QY 3038 GCTGCATCGGTTCTGAGAGAGTGCATCAGAGGTGTGAATCAGACATCCAGTGC 3097  
DB 241 GCTGCATCGGTTCTGAGAGAGTGCATCAGAGGTGTGAATCAGACATCCAGTGC 300  
QY 3098 CCCCTCCCTCAGGAGTATCAGAGGAAAGCCAGTGAACACTAAACAGAGAGACATATG 3157  
DB 301 CCCCTCCCTCAGGAGTATCAGAGGAAAGCCAGTGAACACTAAACAGAGAGACATATG 360  
QY 3158 CACCTCTGACCTTCCCTCCGAGAGCCATCACTCTAATAGAGGAGTGAAGTGCAG 3217  
DB 361 CACCTCTGACCTTCCCTCCGAGAGCCATCACTCTAATAGAGGAGTGAAGTGCAG 420  
QY 3218 GTGGGCTGGGCGCACCAGGAGAGTGAATGCCCTTCTCCCTTCTGAGACACTTCAT 3277  
DB 421 GTGGGCTGGGCGCACCAGGAGAGTGAATGCCCTTCTCCCTTCTGAGACACTTCAT 480  
QY 3278 GTCCCTTCT 3337  
DB 481 GTCCCTTCT 540

QY 3338 TGGGATCCTCTCCACCCCTCTCTAGCCATCCCTTGGGGAAGGTTGGGGAATAATAGCA 3397  
DB 541 TGGGATCCTCTCCACCCCTCTCTAGCCATCCCTTGGGGAAGGTTGGGGAATAATAGCA 600  
QY 3398 TAGACACTGACATGAGCCATTGGAGACACTGGGCCCCCACTGAGACACATGATCTGG 3457  
DB 601 TAGACACTGACATGAGCCATTGGAGACACTGGGCCCCCACTGAGACACATGATCTGG 660  
QY 3458 AG 3459  
DB 661 AG 662

Search completed: October 5, 2002, 16:12:29  
Job time: 8690 sec